

## REMARKS

This is an amendment with request for continued examination filed under 37 C.F.R. 1.114.

### I. Claim Changes

Additional changes have been made in the main independent method claim, which is now amended claim 3, to better and more clearly distinguish the claimed invention from the cited prior art.

The changes in claim 3 should make the main feature of the invention that distinguishes it from the cited prior art somewhat clearer: the phonetic translation hint for a word in the textual description is provided only once in the data stream so that it does not need to be repeated when the word is repeated in the textual description. Instead the same phonetic translation hint is used over and over each time the word is repeated. Since it does not need to be repeated each time the word is repeated, as is the case in the case of prior art embedded pronunciation hints, there is an advantageous reduction in the required number of bits in the data stream. Also the feature that the phonetic translation hint is valid for only a part of the textual description has been deleted from claim 3 to make claim 3 somewhat clearer.

New claim 13 has been added which claims a method including the step

of specifying a part of the textual description in which the phonetic translation hint is valid (see page 4 of applicants' specification), for example by means of XML tags (see page 8 and page 12, last paragraph, of applicants' specification). Then the phonetic translation hint for pronunciation of a repeated word is valid for that part of the textual description and is not repeated at each occurrence of the repeated word. However the phonetic transcription of the repeated word provided by the not-repeated phonetic translation hint is used at every occurrence of the repeated word to define the pronunciation of that repeated word in the data stream.

## **II. Obviousness Rejection based on Huang and Coorman**

Claims 1 to 2, 4 to 6, 8 to 9 and 11 to 12 were rejected under 35 U.S.C. 103 (a) over Huang, et al (referred to below as "Huang") and Coorman, et al, (referred to below as "Coorman").

Claim 3 has been amended so that it is now an independent claim including features and limitations from claims 1 and 2.

Cancellation of claims 1 and 2 and the change in the dependency of the dependent claims so that they all depend on amended independent claim 3 has obviated this rejection of claims 1 and 2 under 35 U.S.C. 103 (a).

### III. Obviousness Rejection based on Huang, Coorman and Carter

Claim 3 was rejected as obvious under 35 U.S.C. 103 (a) over Huang, et al, (referred to below as "Huang") and Coorman, et al, (referred to below as "Coorman"), and further in view of Carter, et al, (referred to below as "Carter").

Huang and Coorman and their relationship to the claimed inventive method have been discussed in the previously filed amendment.

Full credit should be given to the applicants' definition of "multimedia data" on page 5 of the applicants' specification. Applicants are their own lexicographers (M.P.E.P. 2173.05 (a)). The term, "multimedia data", should not be interpreted more broadly than the definition on page 5 of the specification, which is "audio-visual information". The term "multimedia data" in the claims should be accorded this meaning.

According to claim 3 the multimedia data stream includes phonetic translation hints that determine how respective repeated words following the phonetic translation hints are converted or translated into speech by a speech synthesizer. The phonetic translation hints for the respective repeated words associated with them are given only once in the data stream and are not repeated in the textual description at each occurrence of a repeated word in the textual description, as they are in the case of the prior art.

In practice in the method of the invention (claim 3 or 13) predefined (prosodys) information is not used in the decoder nor is a phonetic transcription

repeated after each word in the textual description. Instead in the practice of the methods of claim 3 or 13 a so-called "prosody-lookup-table" for specific words is transmitted, which is used for all repeated words in the following text without the need to repeated any information for decoding.

This feature of the method of claim 3 or 13 provides a savings in processing work because the data stream is shorter, especially if the word or words provided with the phonetic translation hints are repeated a considerable number of times. This provides the flexibility that is needed so that foreign pronunciations of foreign words embedded in a multimedia data stream in a native language that is already recognized by the synthesizer can be properly pronounced, as they would by a foreigner speaking that language (see page 15 of applicants' specification).

Huang and Coorman do not disclose or suggest this distinguishing feature of the method of claim 3, a fact that is recognized in the final Office Action because Carter is needed to allegedly suggest this feature. However applicants respectfully submit that Carter does not suggest all the modifications of the subject matter of Huang and Coorman, which are necessary to arrive at the inventive method as claimed in claim 3.

Carter discloses a method and apparatus and computer program for reducing load on a text-to-speech converter in a message system capable of text-to-speech conversions of E-mail documents (title, abstract) using MPEG-4 TTS. Column 4, lines 1 to 23, does disclose a converter apparatus with a cache memory that stores certain text segments (words) of a received E-mail message

in a cache memory along with the converted speech signals for those text segments. Then when playback of other E-mail messages is requested if the other E-mail messages contained a text segment (word) that is stored in the cache memory along with the corresponding speech signals, the text-to-speech conversion process for that text segment (word) is by-passed and instead the stored speech signal is retrieved. For example, see figure 3 and the description associated with it, especially column 3, line 67 to column 4, line 10, which states:

"Upon a request by a user to convert the text segments of a new e-mail message to speech signals for playback via a telephone handset 34, a controller 42 determines whether any of the text segments of the new e-mail message are identical to previously converted text segments for which speech signals are already stored in the cache 40. If so, the stored speech signals of those text segments are played back to the user from the cache, thus avoiding the need for the text-to-speech converter to convert those test segments of the new e-mail message to speech."

There are advantages to associating a respective phonetic translation hint with a particular word or text segment during a portion of the data stream instead of storing the speech signals associated with the word, although then it is necessary to repeat the speech synthesis or conversion process each time the word or text segment is repeated. However in the method of claim 3 each time the word is repeated it is not immediately preceded by the associated phonetic translation hint. This significantly shortens the data stream in the case of words which are repeated a large number of times and provides the flexibility to cope with special cases where automatic transcription is not applicable (see applicants' specification on page 3, lines 13 to 16). However the speech

conversion process is repeated each time a repeated word occurs but with the same phonetic transcription that has been specified and indicated as valid for the entire textual description or specified part of it.

Carter clearly does not suggest associating a particular phonetic translation hint with a given repeated word over a specified portion of a textual description as claimed in claim 13. Instead actual speech signals are associated with the given word. These actual speech signals are stored in memory so that they are retrieved and played if the word is repeated, as shown by the quotation from column 4 of this reference.

However the method of Carter is inefficient and consumes much memory or requires longer time intervals, because of the complexity of digital or analog speech signals necessary for faithful sound reproduction, since large amounts of data must be retrieved and/or stored.

Thus Carter only suggests storing speech signals or associating a respective speech signal with a particular word so that repetition of the conversion process is unnecessary when the word is repeated. Carter does not disclose or suggest associating a respective phonetic translation hint with a particular word, so that the translation hint does not need to be repeated, as claimed in claim 3. These are entirely different methods and the method of Carter would not suggest the method of claim 3. Carter does not suggest all the modifications of the subject matter of Huang or Coorman necessary to arrive at the method as claimed in claim 3.

Also, the method of applicants' claim 3 requires that the previous phonetic

transcription hint is valid for a defined portion or all of the textual description in the data stream (claim 3). No storage and no retrieval of speech signals or any other data in and from a memory are necessarily required in the claimed method of applicants' claim 3 in contrast to the method of claim 5 of Carter, which is described in part in column 4. For example, the data stream could include control characters (see page 11 and 13 of applicants' specification) that define the portion of the data stream over which repetition of the hint is not necessary so that it would not be necessary to store the hint itself.

In addition, because storage and retrieval of speech signals in a memory is required in the method of Carter, the savings due to the reduction in word-to-speech signal conversion work when text segments are repeated is counteracted by the time required for storage and retrieval of the speech signals. The embodiments using the cache memory (for speed) are particularly limited, as explained in column 4, lines 13 to 23, of Carter, because the cache can only function with repeated text segments that have 40 or fewer characters. There are no such limits due to storage and retrieval work that are necessarily present in applicants' method of claim 3. The phonetic translation hint could in principle be larger than 40 characters and apply to a large text segment, for example, an entire sentence or phrase in a foreign language.

It is well established by many U. S. Court decisions that to reject a claimed invention under 35 U.S.C. 103 there must be some hint or suggestion in the prior art of the modifications of the disclosure in a prior art reference or references used to reject the claimed invention, which are necessary to arrive at

the claimed invention. For example, the Court of Appeals for the Federal Circuit has said:

"Rather, to establish obviousness based on a combination of elements disclosed in the prior art, there must be some motivation, suggestion or teaching of the desirability of making the specific combination that was made by the applicant...Even when obviousness is based on a single reference there must be a showing of a suggestion of motivation to modify the teachings of that reference." *In re Kotzab*, 55 U.S.P.Q. 2<sup>nd</sup> 1313 (Fed. Cir. 2000). See also M.P.E.P. 2141

The term "phonetic translation hint" cannot be reasonably interpreted as including "speech signals". Speech signals are the resulting signals output from a speech synthesizer conversion or translation process. However "phonetic translation hints" are data, which is input to a speech synthesizer or converter to produce the resulting speech signals. An output signal from a special processor cannot be held to be the same as the input signal.

Thus Carter does not reasonably suggest storing and retrieving phonetic translation hints for a particular word in a portion of a data stream so that repetition of the translation hint is not required when the word is repeated. Carter would only suggest storing the resulting speech signals of a word so that the conversion process does not need to be repeated or associating the converted speech signals with a particular word so that the conversion is not necessary when the word is repeated. The one method is not obvious from the other.

For the foregoing reasons and because of the changes in claim 3, withdrawal of the rejection of claim 3 as obvious under 35 U.S.C. 103 (a) over Huang, et al, and Coorman, et al, and further in view of Carter, et al, is respectfully requested.

Similarly for the same reasons amended dependent claims 4 to 6, 8, 9, 11 and 12 should not be rejected under 35 U.S.C. 103 (a) over Huang, et al, and Coorman, et al, and further in view of Carter, et al.

In addition it is respectfully submitted that new claim 13 should not be rejected under 35 U.S.C. 103 (a) over Huang, et al, and Coorman, et al, and further in view of Carter, et al.

#### **IV. Obviousness Rejection based on Huang, Coorman and Sharman**

Claims 7 and 10 were rejected as obvious under 35 U.S.C. 103 (a) over Huang, et al, (referred to below as "Huang") and Coormann, et al, (referred to below as "Coormann"), and further in view of Sharman, et al (referred to below as "Sharman").

The features of claims 7 and 10 are currently not being relied on to establish patentability of the claimed method, but the dependent claims have been amended so that they now depend on claim 3. Instead these features are features of preferred embodiments of the amended method claim 3.

For the foregoing reasons withdrawal of the rejection of claims 7 and 10 as obvious under 35 U.S.C. 103 (a) over Huang, et al. and Coorman, et al, and further in view of Sharman, et al, is respectfully requested.

Should the Examiner require or consider it advisable that the specification, claims and/or drawing be further amended or corrected in formal respects to put this case in condition for final allowance, then it is requested that such amendments or corrections be carried out by Examiner's Amendment and the case passed to issue. Alternatively, should the Examiner feel that a personal discussion might be helpful in advancing the case to allowance, he or she is invited to telephone the undersigned at 1-631-549 4700.

In view of the foregoing, favorable allowance is respectfully solicited.

Respectfully submitted,



Michael J. Striker,  
Attorney for the Applicants  
Reg. No. 27,233